



What is Epidemiology?

Epidemiology

- **Epidemiology is the study of the determinants, distribution, and frequency of disease**
- **Who gets disease and why**
- **Epidemiologists study sick and well people to determine the crucial difference between those who get disease and those who are spared**

Important Issues that Epidemiology can Address

Disease

Mortality

Hospitalization

Disability

Quality of Life

Health Status

Epidemiology

- **Epidemiology weighs and balances**
- **Epidemiology contrasts and compares**
- **Epidemiologists use RATES**
 - **events/population at risk**

Epidemiology

- **Numerator**
 - the number of people to whom something happened (i.e. they got sick, died, etc.)
- **Denominator**
 - the population at risk -- all the people at risk for the event

What is Epidemiology?

- **Some describe it as the study of epidemics**
- **What is an epidemic?**
 - **An epidemic occurs when there are significantly more cases of the same disease than past experience would have predicted.**

Uses of Epidemiology

- **To study the cause (or etiology) of disease(s), or conditions, disorders, disabilities, etc.**
 - **determine the primary agent responsible or ascertain causative factors**
 - **determine the characteristics of the agent or causative factors**
 - **define the mode of transmission**
 - **determine contributing factors**
 - **identify and determine geographic patterns**

Uses of Epidemiology

- **To determine, describe, and report on the natural course of disease, disability, injury, and death.**
- **To aid in the planning and development of health services and programs**
- **To provide administrative and planning data**

Example: Use of Epidemiology

- In the United States, the National Center for Health Statistics is a data source for information on health and disease
 - <http://www.cdc.gov/nchs>
- Globally, the World Health Organization is a data source for information on health and disease
 - <http://www.who.int>

Leading Causes of Death, USA, 2000

(number of deaths)

1. Heart Disease
2. Neoplasms
3. Cerebrovascular Disease
4. Chronic Pulmonary Disease
5. Accidents/Injuries
6. Diabetes mellitus
7. Influenza and pneumonia
8. Alzheimer's Disease
9. Nephritis, nephrotic syndrome
10. Septicemia

Leading Causes of Death Worldwide, 2000

(Based on number of global deaths)

1. Ischemic Heart Disease
2. Cerebrovascular Disease
3. Lower Respiratory Infections
4. HIV/AIDS
5. COPD
6. Perinatal Conditions
7. Diarrhoeal Diseases
8. Tuberculosis
9. Road Traffic Injuries
10. Lung Cancers

Pakistan



Pakistan: Leading Causes of Premature Death, 1990

1. Diarrhea
2. Lower Respiratory Infections - child
3. Tuberculosis
4. Rheumatic heart disease
5. Chronic liver disease
6. Congenital malformations
7. Birth Diseases
8. Ischemic heart disease
9. Child Septicemia
10. Injuries

Purpose of Epidemiology

- To provide a basis for developing disease control and prevention measures for groups at risk. This translates into developing measures to prevent or control disease.

Two Broad Types of Epidemiology

- **Descriptive Epidemiology**
 - Examining the distribution of disease in a population, and observing the basic features of its distribution
- **Analytic Epidemiology**
 - Testing a hypothesis about the cause of disease by studying how exposures relate to the disease

Descriptive Epidemiology is the Antecedent to Analytical Epidemiology

- **Analytic epidemiology studies
require information to**
 - **know where to look**
 - **know what to control for**
 - **develop viable hypotheses**

Three essential characteristics of disease that we look for in descriptive studies are...

- Person**
- Place**
- Time**

Person

- **Age, gender, ethnicity**
- **Genetic predisposition**
- **Concurrent disease**
- **Diet, exercise, smoking**
- **Risk taking behavior**
- **Education, occupation**

Place

- **Geographic place**
 - presence or agents or vectors
 - climate
 - geology
 - population density
 - economic development
 - nutritional practices
 - medical practices

Time

- **Calendar Time**
- **Time since an event**
- **Physiologic cycles**
- **Age (time since birth)**
- **Seasonality**
- **Temporal trends**

Example

- **You have been asked to investigate an event in which 2,220 people were exposed and 1,520 of them died.**
- **Your role as an epidemiologist is to ask questions about person, place and time.**

How do we ask questions?

Surveys

-of survivors

-of next-of-kin

-of other related persons

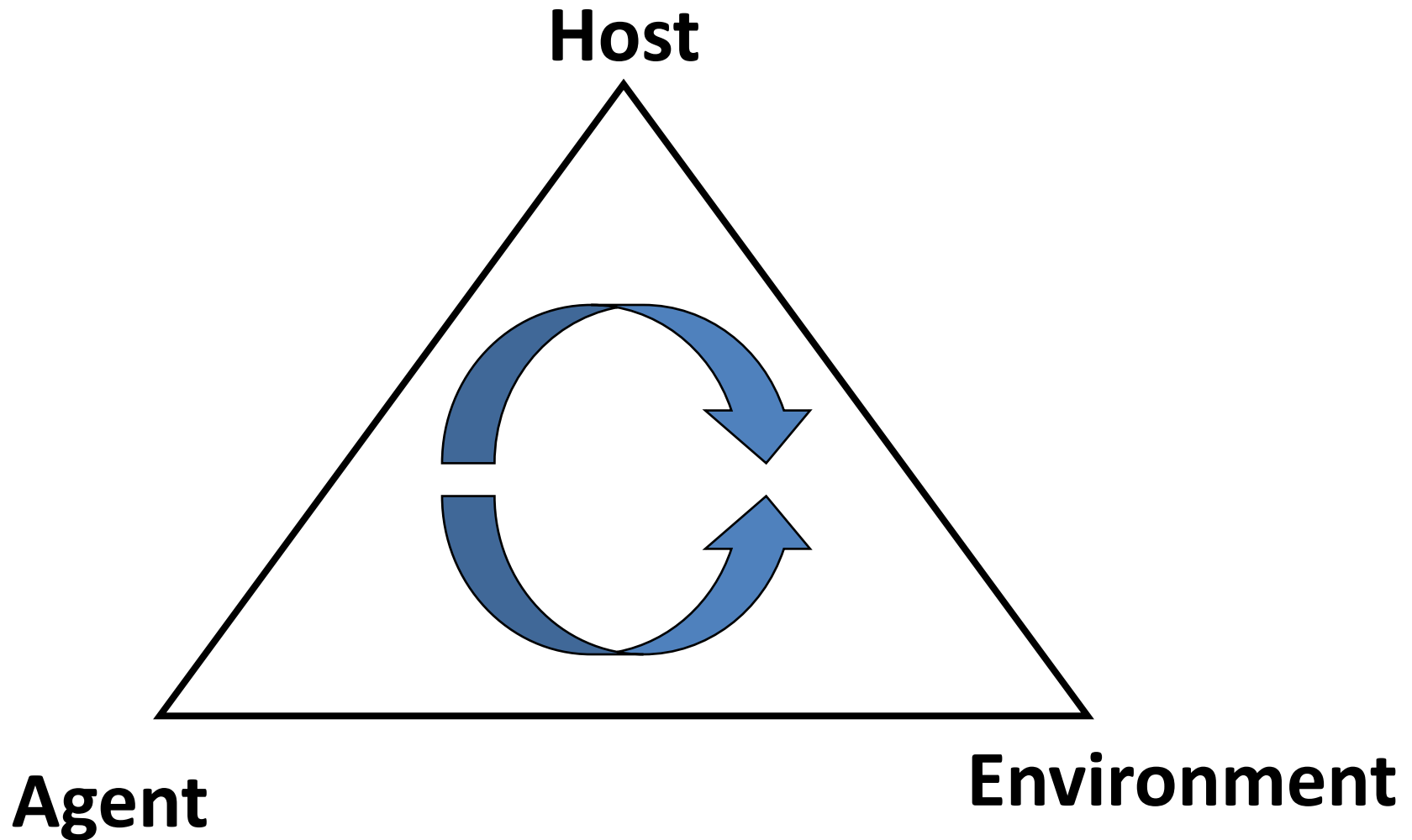
with questions you learn that ...

- **Person: Men, women and children were all exposed and at risk. The majority of people who died were wealthy and young men between 18-50 years (when compared to survivors).**
- **Place: All those exposed were within 1 block of one another, the climate was cold.**
- **Time: Mid April, people died within hours of the precipitating exposure.**

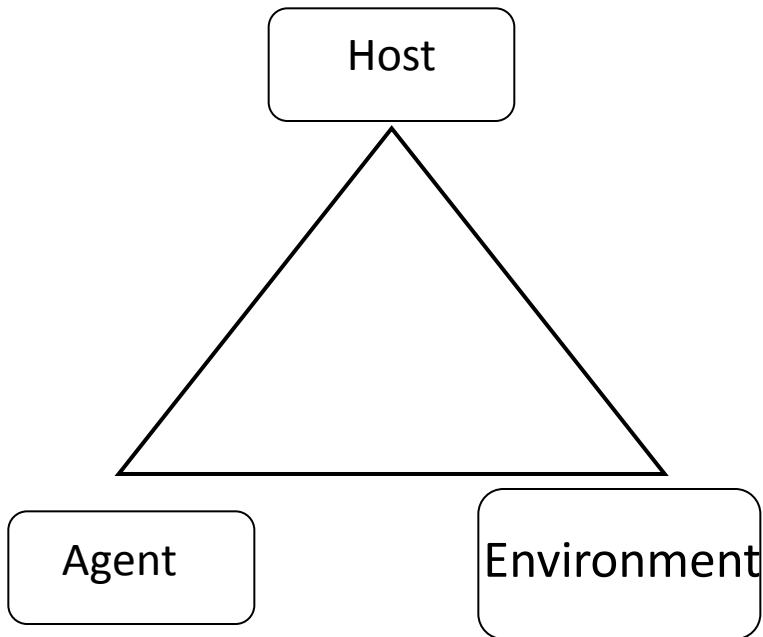
Three essential characteristics that are examined to study the cause(s) for disease in analytic epidemiology are...

- Host**
- Agent**
- Environment**

Epidemiologic Homeostasis



The Epidemiologic Triangle



Host Factors

Personal traits

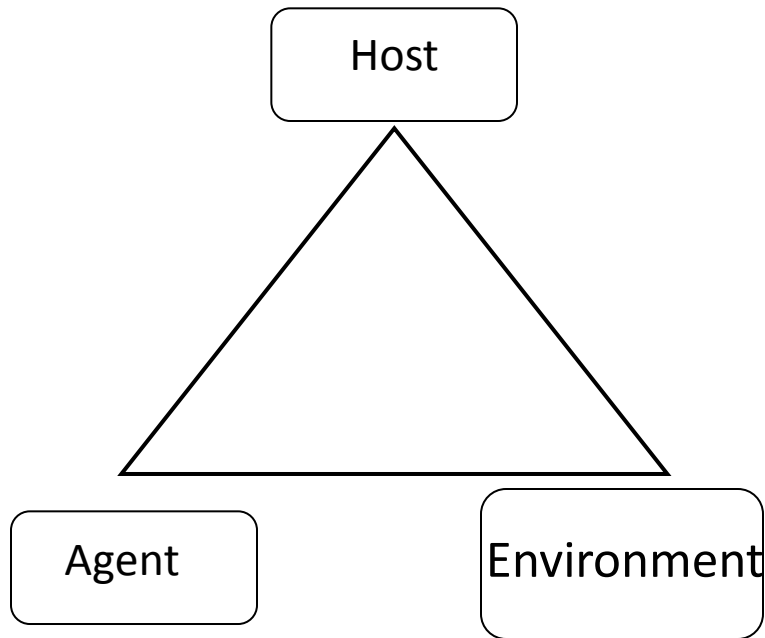
behaviors

genetic predisposition

immunologic factors

- Influence the chance for disease or its severity

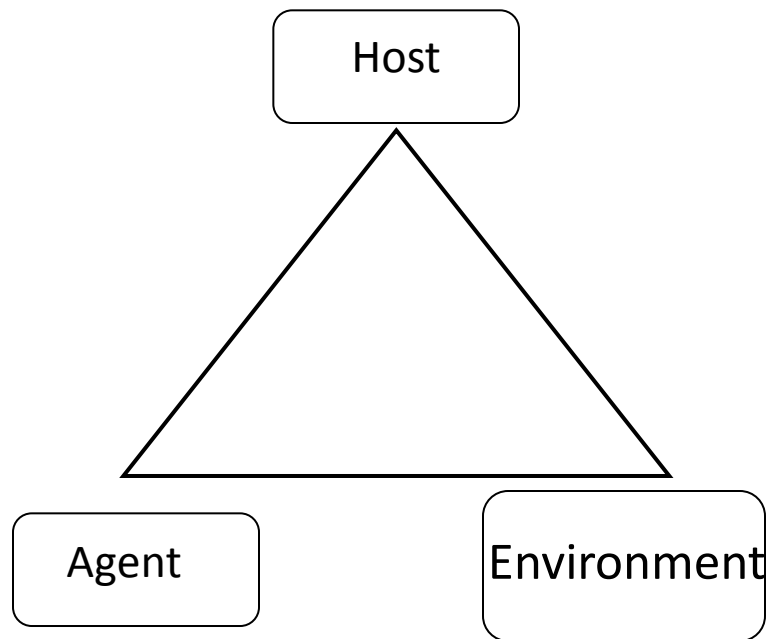
The Epidemiologic Triangle



Agents
Biological
Physical
Chemical

- **Necessary for disease to occur**

The Epidemiologic Triangle



Environment

External conditions
Physical or biologic
or social

- Contribute to the disease process

Epidemics arise when host, agent, and environmental factors are not in balance

- **Due to new agent**
- **Due to change in existing agent (infectivity, pathogenicity, virulence)**
- **Due to change in number of susceptibles in the population**
- **Due to environmental changes that affect transmission of the agent or growth of the agent**

Epidemiologic Activities

...are often framed under the mantle of descriptive and analytic epidemiology

- Descriptive epidemiology – person, place & time

- Demographic distribution
- Geographic distribution
- Seasonal patterns etc.
- Frequency of disease patterns

- Useful for:

- Allocating resources
- Planning programs
- Hypotheses development



Epidemiologic Activities

- **Analytic epidemiology**
 - built around the analysis of the relationship between two items
 - Exposures
 - Effects (disease)
 - looking for determinants or possible causes of disease
 - useful for
 - hypothesis testing



Epidemiology

Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to control health problems.

Example

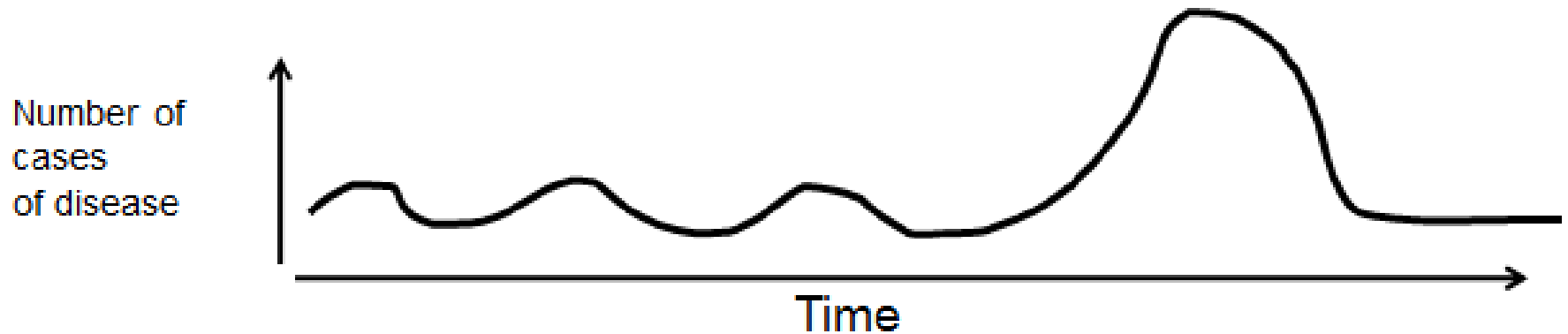
- What is the incidence of Type II diabetes between 2010-2015 among overweight subjects born in 1970?
- What are the possible causes of Type II diabetes in this population?
- If we were to change population dietary habits, what improvement in the incidence could we affect?

- **Determinant:** any factor, whether event, characteristic, or other definable entity, that brings about a change (negative or positive) in a health condition or other defined characteristic.
- **Epidemiology** is also used to search for determinants, which are the causes and other factors that influence the occurrence of disease and other health-related events.
- **Epidemiologists** assume that illness does not occur randomly in a population, but happens only when the right accumulation of risk factors or determinants exists in an individual.
- E.g.. 1. Is smoking associated with higher risk of lung cancer. Is obesity associated with higher risk of Type II diabetes. Is high fiber food associated with lower risk of colon cancer

Endemic - The habitual presence (or usual occurrence) of a disease within a given geographic area

Epidemic - The occurrence of an infectious disease clearly in excess of normal expectancy, and generated from a common or propagated source

Pandemic - A worldwide epidemic affecting an exceptionally high proportion of the global population



Morbidity vs. Mortality

Morbidity is any physical or psychological state considered to be outside the realm of normal well-being. The term morbidity is often used to describe illness, impairment, or degradation of health, especially when discussing chronic and [age-related diseases](#) which can worsen over time. The higher your morbidity, the shorter your [lifespan](#) may be that if you were healthy.

While morbidity refers to your level of health and well-being, mortality is related to your risk of death. They are not the same thing. Morbidity doesn't necessarily mean that your ill-health is immediately life-threatening. Over time, however, if an illness continues it may increase your risk of mortality (death). Current research shows people are now living longer with diseases than they once were

Prevalence vs. Incidence

- Prevalence is the number of existing cases of disease in the population during a defined period.

$$\frac{\text{\# of Cases Existing in a Given Population at a Single Point in Time}}{\text{Population at That Time}}$$

- Incidence is the number of new cases of disease that develop in the population **at Risk of developing a disease** during a defined period.

$$\frac{\text{\# of New Cases Occurring in a Given Population in a Specified Time Period}}{\text{Population in That Time Period}}$$

Prevalence vs. Incidence

- The easy way to remember the **difference** is that **prevalence** is the proportion of cases **in the** population at a given time rather than rate of occurrence of new cases. Thus, **incidence** conveys information about the risk of contracting the disease, whereas **prevalence** indicates how widespread the disease is.

Prevalence vs. Incidence

- Consider the following example. Say you are looking at a sample population of 225 people, and want to determine the incidence rate of developing HIV over a 10-year period:
- At the beginning of the study ($t=0$) you find 25 cases of existing HIV. These people are not counted as they cannot develop HIV a second time.
- A follow-up at 5 years ($t=5$ years) finds 20 new cases of HIV.
- A second follow-up at the end of the study ($t=10$ years) finds 30 new cases.

Prevalence vs. Incidence

- If you were to measure prevalence you would simply take the total number of cases ($25 + 20 + 30 = 75$) and divide by your sample population (225).
- So prevalence would be $75/225 = 0.33$ or 33% (by the end of the study). This tells you **how widespread** HIV is in your sample population, but little about the actual risk of developing HIV for any person over a coming year.

Prevalence vs. Incidence

- To measure incidence you must take into account how many years each person contributed to the study, and when they developed HIV.
- When it is not known exactly when a person develops the disease in question, epidemiologists frequently use the actuarial method, and assume it was developed at a half-way point between follow-ups.
- **In this calculation:**
 - At 5 yrs you found 20 new cases, so you assume they developed HIV at 2.5 years, thus contributing $(20 * 2.5) = 50$ person-years of disease-free life.

Prevalence vs. Incidence

- At 10 years you found 30 new cases. These people did not have HIV at 5 years, but did at 10, so you assume they were infected at 7.5 years, thus contributing $(30 * 7.5) = 225$ person-years of disease-free life.

That is a total of $(225 + 50) = 275$ person years so far.

- You also want to account for the 150 people who never had or developed HIV over the 10-year period, $(150 * 10)$ contributing 1500 person-years of disease-free life.
- That is a total of $(1500 + 275) = 1775$ person-years of life. Now take the 50 new cases of HIV, and divide by 1775 to get 0.028, or 28 cases of HIV per 1000 population, per year. In other words, if you were to follow 1000 people for one year, you would see 28 new cases of HIV.

Retrospective study: looks backwards and examines exposures to suspected risk or protection factors in relation to an outcome that is established at the start of the study.

Prospective study: watches for outcomes, such as the development of a disease, during the study period and relates this to other factors such as suspected risk or protection factor(s). The study usually involves taking a cohort (group) of subjects and watching them over a long period.

Basic Measures of Association: In the above studies, we often need to know the relationship between an outcome and certain factors (e.g., age, sex, race, smoking status, etc.). Relative risk & odds ratio (explained next slides) are used for this purpose.

Case-Control Studies

Case-control studies select subjects based on their disease status.

- It is a **retrospective study**.
- A group of individuals that are disease positive (the "case" group) is compared with a group of disease negative individuals (the "control" group).
- The control group should ideally come from the same population that gave rise to the cases.
- The case-control study looks back through time at potential exposures that both groups (cases and controls) may have encountered.
- Case control studies are observational because no intervention is attempted and no attempt is made to alter the course of the disease.

A 2×2 table is constructed, displaying exposed cases (A), exposed controls (B), unexposed cases (C) and unexposed controls (D).

Contingency (or 2 x 2) Table

	Cases	Controls	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

$$\text{OR} = (a/c) / (b/d)$$
$$= (a*d) / (b*c)$$

The **odds** are defined as the **probability** that the event (exposed) will occur divided by the **probability** that the event will not occur (unexposed).

Odds = $P/(1-P)$. For cases = $(a/(a+c))/(c/a+c) = a/c$

Similarly for controls Odds = b/d

The statistic generated to measure association is the **odds ratio (OR)**, which is the ratio of the odds of exposure in the cases (A/C) to the odds of exposure in the controls (B/D), i.e. **OR = (AD/BC)**.

Case-control studies: Example

Suppose that unusual number of patients (19) diagnosed with salmonella were admitted to a hospital on one day, and all patients were from confined geographic area. A sample of 38 non diseased people as a comparison group (the controls). In this case, the "controls" were non diseased people who were matched to the cases with respect to age, gender, and neighborhood of residence. The case and control groups answered a questioner listing several types of food that may be sources of infection. The results of two types of food are listed in next slide.

OR of 1 suggests that there is no difference between the groups; i.e. there would be no association between the suggested exposure (fish) and the outcome (being ill)

OR of > 1 suggests that the odds of exposure are positively associated with the adverse outcome compared to the odds of not being exposed

OR of < 1 suggests that the odds of exposure are negatively associated with the adverse outcomes compared to the odds of not being exposed. Potentially, there could be a protective effect

	Cases	Controls
Ate eggs	10	19
Did not eat eggs	9	19
Total	19	38

Odds Ratio = $(10/9) / (19/19) = 1.1$
(close to one). This certainly provides no compelling evidence to suggest an association the cases and egg consumption.

	Cases	Controls
Drank milk	18	7
Did not drink milk	1	29
Total	19	38

Odds Ratio = $(18/1) / (7/29) = 75$
This suggests who drank milk had 75 times the risk of being a case compared to those who did not.

Cohort studies

- Cohort studies select subjects based on their exposure status. The study subjects should be at risk of the outcome under investigation at the beginning of the cohort study; this usually means that they should be disease free when the cohort study starts.
- The cohort is followed through time to assess their later outcome status.
- An example of a cohort study would be the investigation of a cohort of smokers and non-smokers over time to estimate the incidence of lung cancer. The same 2×2 table (shown next slide) is constructed as with the case control study.
- However, the point estimate generated is the relative risk (RR), which is the probability or incidence of disease for a person in the exposed group, $P = A / (A + B)$ over the probability or incidence of disease for a person in the unexposed group, $P = C / (C + D)$.

Calculation of Relative Risk

	Outcome		
Exposure	Case	Absent	TOTAL
Exposed	a	b	a+b
Unexposed	c	d	c+d
TOTAL	a+c	b+d	<i>a+b+c+d</i>

$$\text{Relative Risk} = \frac{\left(\frac{a}{a+b} \right)}{\left(\frac{c}{c+d} \right)}$$

Relative Risk

- **Ratio of the incidence rates between two groups**
- **Can only be calculated from prospective studies (cohort studies)**
- **Interpretation**
 - **$RR > 1$: Increased risk of outcome among “exposed” group**
 - **$RR < 1$: Decreased risk, or protective effects, among “exposed” group**
 - **$RR = 1$: No association between exposure and outcome**

Calculation of Relative Risk

$$RR = \frac{\text{incidence or probability among exposed}}{\text{incidence or probability among non-exposed}}$$

Relative Risk Case Study

Smoking and low birth weight

	Birth Weight		
Smoking status	<2500 g	≥2500 g	TOTAL
Smoker	120	240	360
Non-smoker	60	580	640
TOTAL	180	820	<i>1000</i>

Answers to Relative Risk Case Study

1. Incidence of LBW among smokers

$$= \frac{120}{360} = 0.333$$

2. Incidence of LBW among non-smokers

$$= \frac{60}{640} = 0.0938$$

3. **Relative risk** for having a LBW baby among smokers versus non-smokers

$$= \frac{0.333}{0.0938} \approx 3.6$$

The study **suggests**, the relative risk of having LBW baby among smoker women is 3.6 times that of non-smoker women.